Exercise capacity in non-specific chronic low back pain patients

Hodselmans, Audy P.; Dijkstra, Pieter U.; Geertzen, Jan H. B.; van der Schans, Cees P.

Published in:
Journal of Occupational Rehabilitation

DOI:
10.1007/s10926-008-9143-2

Publication date:
2008

J Occup Rehabil (2008) 18:282–289
DOI 10.1007/s10926-008-9143-2

Exercise Capacity in Non-Specific Chronic Low Back Pain Patients: A Lean Body Mass-Based Åstrand Bicycle Test; Reliability, Validity and Feasibility

Audy P. Hodselmans · Pieter U. Dijkstra · Jan H. B. Geertzen · Cees P. van der Schans

Published online: 15 July 2008

Abstract *Objective* Measurement of exercise capacity is essential in patients with non-specific chronic low back pain (CLBP). However, the conventional Åstrand bicycle test is not feasible in patients with a very poor aerobic capacity. Therefore the Åstrand bicycles test for non-specific CLBP patients based on lean body mass (LBM) was developed as an alternative. The aim of this study was to evaluate reliability and validity of the LBM-based Åstrand test. *Subjects* Twenty patients with non-specific CLBP and 20 healthy subjects were included for the reliability evaluation, and 19 healthy subjects for the validity evaluation. *Method* Patients and healthy subjects were assessed twice. Intra class correlation (ICC), repeatability coefficient (RC) and the limits of agreement (LOA) were calculated as a measure of test re-tests reliability. An ICC $\geq$ 0.75 was considered acceptable. Validity was tested by calculating ICC between the LBM-based Åstrand test and a maximal bicycle test. *Results* The LBM-based Åstrand test shows good reliability, reflected by an ICC $\geq$ 0.91 and 95% of the 20 patients could perform the test. However, differences with the estimated true value reflected by the RC and natural variation reflected by the LOA were substantial in patients. Validity was good, reflected by ICC $\geq$ 0.88. *Conclusion* The present study shows that the LBM-based Åstrand test is a reliable, valid, and feasible method for patients with non-specific CLBP. However, a substantial amount of variation should be taken into account in patients when interpreting the test results clinically.

Keywords Exercise capacity · Reproducibility of results · Chronic low back pain · Åstrand test

A. P. Hodselmans (✉) · C. P. van der Schans
Center for Applied Research and Innovation in Health Care and in Nursing, Hanze University Groningen Applied Sciences, Eyssoniusplein 18, 9714 CE Groningen, The Netherlands
e-mail: A.P.Hodselmans@pl.hanze.nl

A. P. Hodselmans · P. U. Dijkstra · J. H. B. Geertzen
Graduate School for Health Research, University of Groningen, A. Deusinglaan 1, 9713 AV Groningen, The Netherlands

P. U. Dijkstra · J. H. B. Geertzen
Center for Rehabilitation, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands

P. U. Dijkstra
Department of Oral and Maxillofacial Surgery, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands

Introduction

Deconditioning in terms of exercise physiology is associated with a loss of exercise capacity (VO_2 max) [1]. In patients with non-specific chronic low back pain (CLBP) deconditioning is thought to be both a cause and a consequence of non-specific CLBP [1, 2]. Deconditioning, as a factor contributing to the chronicity of non-specific CLBP, forms the basis of aerobic exercise training in CLBP rehabilitation programmes [1–4]. Recently, activity avoidance has been presented as one of the factors perpetuating chronic pain in the fear-avoidance model [5–8]. It is assumed that fear and the feelings of disability may both contribute to the avoidance of activity leading to reduction of aerobic capacity [9]. However, this assumption could not be confirmed [5]. Additionally levels of aerobic fitness in patients with CLBP are comparable with those in healthy subjects [1]. Generally aerobic capacity is measured using a maximal aerobic capacity test; however, this test is strongly influenced by motivation, fear, and pain in patients with non-specific CLBP and is invalid when fear and pain

expectation rather than aerobic capacity limits performance [7, 10, 11]. It has been reported that 54% of the CLBP patients who underwent maximal treadmill testing did not reach maximal performance criteria due to pain [7, 12].

Submaximal tests, like the Åstrand test, are used as an alternative for maximal exercise testing. The reliability of the submaximal Åstrand bicycle test is good in healthy subjects [7, 12]. However, problems are expected when using this test in patients with a very poor aerobic capacity like patients with non-specific CLBP because workload is intensively increased during the first 2 min without taking into account the individual aerobic capacity [5, 7]. Clinically, many non-specific CLBP patients are unable to finish the Åstrand bicycle test because the initial workload was set too high [7]. Therefore, estimation of the maximum oxygen uptake from the submaximal test results may not be possible in a substantial number of patients. In the Siconolfi protocol (1982), which is specifically designed to assess aerobic performance in healthy inactive persons, load is increased more gradually, and the initial load is lower, compared with the Åstrand test [5]. This protocol starts with 25 W and increases with 25 W every 2 min until the subject achieved target heart rate [5]. The disadvantage of this protocol is the long duration some subjects have to cycle to achieve the target heart rate. This problem can be avoided if the individual workload is tailored on lean body mass (LBM). LBM decreases over 30 days of rest while body weight does not. This phenomenon suggests that LBM reflects the state of loss of muscle mass and deconditioning, related to avoidance of activity in non-specific CLBP patients [6, 13]. A LBM-based Åstrand test in a subject with a poor aerobic capacity has smaller increases in workload as compared to a subject with a very good aerobic capacity. In order to decrease the test duration and to increase the number of patients that can finish the bicycle test, an LBM-based Åstrand bicycle test was developed where the predefined workload increase schedule is based on LBM [7, 9, 14].

The purpose of this study was to determine the reliability, validity, and feasibility of the LBM-based Åstrand bicycle test.

Methods

Participants

The test–retest reliability of the LBM-based Åstrand bicycle test was evaluated in 20 patients with non-specific CLBP (12 women) and 20 healthy pain-free subjects, group I (10 women) (Table 1). Patients were recruited from the Center for Rehabilitation at the University Medical Center Groningen, The Netherlands. Healthy subjects were recruited from the student body of the Institute for Human Movement Sciences of the University of Groningen, The Netherlands.

The validity of the LBM-based Åstrand bicycle test was evaluated in another group of 19 healthy pain-free subjects, group II (11 women). These healthy subjects were also students from the Institute for Human Movement Sciences of the University of Groningen, The Netherlands.

Procedure

Admission assessment of patients was carried out by a rehabilitation physician at the Center for Rehabilitation from the University Medical Center Groningen before patients entered the study. Twenty patients diagnosed with non-specific CLBP who participated in a rehabilitation program were included in the study if they were between 18 and 65 years of age. The rehabilitation physician used general admittance criteria: 1) non-specific CLBP (i.e., pain lasting for more than 3 months, LBP without shown organic substratum); and 2) patients were content with the diagnostic process and motivated for the treatment program. Patients were excluded if they were in: 1) a conflict situation with employer or insurance company regarding their work; 2) a financially profitable situation caused by their illness; and 3) specific low back pathology, co-morbidity, pregnancy, and psychopathology. Any medical condition that could interfere with physical performance tests, major surgery within the previous year, existing infectious disease, cancer, neuralgic, or cardiovascular disease were exclusion criteria for both patients and controls. Additionally, healthy subjects were excluded if they had a history of LBP which had lasted more than 1 week, required medical attention, or resulted in absence from work or school within the previous 6 months. The median duration of LBP complaints in patients is 68 months with a range from 8 to 180 months. The mean (SD) score of the Roland Morris Disability Questionnaire is 10.2 (5.3). All participants signed informed consent forms and were assessed in two sessions. Patients were assessed twice, during the waiting time, before starting the cognitive somatic back school rehabilitation program. The second session was performed after a 2-week interval. Time of day, day of week, and place of assessment were kept constant for the two sessions. The aim of this cognitive somatic back school program is to achieve an optimal functional capacity. The program is given by a physical therapist and consists of the following parts: (1) self-treatment according to the principles of McKenzie; (2) individual circuit training, aqua jogging, and sports activities in groups; (3) education concerning overload mechanisms and influence of psychosocial factors on functional capacity and perceived disability.

Table 1 Basic characteristics of the participants, patients ($n = 20$) and healthy subjects group I for the reliability test ($n = 20$) and healthy subjects group II for the validity test ($n = 19$)

SD: standard deviation

	Patients ($n = 20$) mean (SD)	Healthy subjects ($n = 20$) mean (SD) Group I	Healthy subjects ($n = 19$) mean (SD) Group II
Age (year)	33.8 (8.6)	22.0 (1.6)	22.9 (2.2)
Height meters	1.76 (0.1)	1.79 (0.1)	1.74 (0.2)
Weight, kg,	73.9 (14.7)	72.4 (8.5)	70.9 (8.5)
LB M kg	55.2 (10.7)	60.4 (9.7)	57.1 (9.6)
Women (%)	12 (60%)	10 (50%)	11 (58%)

To evaluate validity, 19 healthy subjects, group II was assessed, using the LBM-based Åstrand bicycle test, and the maximal aerobic bicycle capacity test with a 2-week interval.

Measurements

LBM-based Åstrand bicycle test An LBM-based Åstrand bicycle test was performed to estimate the maximum oxygen consumption; VO_2 max in l/min, ml/min/kg Body Mass (BM) and ml/min kg LBM [7, 9, 14]. First LBM was measured according to the Durnin and Womersly protocol using a skinfold calliper (Servier Nederland B.V., Leiden, The Netherlands) [7, 9, 15]. The subjects performed the test on a calibrated Cycle ergometer (Excalibur Sport, Lode B.V., Groningen, The Netherlands). Heart rate (HR) was recorded using a monitor connected to electrodes on the patient's chest (Polar Favour, Kempele, Finland). The subjects started cycling under a predetermined workload of 0.5 W/kg LBM at a constant rate of 60 rates per minute. After 2 min cycling the workload was increased to 1.5 W/kg LBM. If the HR remained below 120 beats/min the workload was increased by 0.5 W/kg LBM every 2 min. Once HR exceeded 120 beats/min, the patient cycled 6 min under a fixed workload to reach a steady state phase, meaning that HR did not vary more than ±5 beats/min during the final 2 min of exercise. The mean HR during the final 2 min of exercise was calculated. The maximum oxygen uptake (VO_2 max) was estimated using the Binkhorst calculation based on the linear association between HR and increase in oxygen uptake, for men and women [9, 14, 16].

$$VO_2\,\max\,(\text{men}) = \frac{174.2 * \text{load Watts} + 4020}{103.2 * \text{HR} - 6299},$$

$$VO_2\,\max\,(\text{women}) = \frac{163.8 * \text{load Watts} + 3780}{104.4 * \text{HR} - 7514}$$

The calculated VO_2 max was corrected for age using an age correction factor from Åstrand [17]. The test was terminated if the subject did not attain a HR of at least 120 beats/min, if the HR exceeded the predetermined maximum (([220 − age]) * 0.85), the systolic/diastolic blood pressure reached a level of 220/115 mm Hg, or if the subject showed signs of serious cardiovascular or pulmonary difficulties. After 6 min cycling under a fixed workload, the load decreased over 1 min to 0.25 W/kg LBM and the subject cycled for 1 min under this workload of 0.25 W/kg LBM.

Maximal Bicycle Test

The maximal bicycle test was performed on a calibrated cycle ergo meter (Excalibur Sport, Lode B.V., Groningen, The Netherlands). During the maximal exercise test the participants breathed through a facemask (Hans Rudolph Inc, USA) connected to a calibrated metabolic cart (Oxygen Champion, Jaeger, Mijnhardt, Bunnink, The Netherlands). Expired gas was passed through a flow meter, oxygen (O_2) analyzer, and a carbon dioxide (CO_2) analyzer. These analyzers were connected to a computer, which calculated breath-by-breath minute ventilation, oxygen consumption (VO_2), CO_2 production, and the respiratory exchange ratio using conventional equations. HR was measured continuously during the test on a bipolar electrocardiogram. The HR was recorded by a monitor connected to electro cardio gram electrodes (3 M red dot) on the patient's chest. Maximal effort was registered when one of two criteria were met, HR > 220 − age + 10 or respiratory exchange ratio >1.0 [18]. The healthy subjects started cycling the first 2.5 min under a predetermined workload of 0.5 W/kg Body Mass (BM) and the second 2.5 min under a workload of 1 W/kg BM. After these 5 min the load increased by 0.25 W/kg BM per min until maximum effort was reached. During the recovery period of 5 min the subjects cycled under a workload of 50 W.

Data Analysis

Descriptive statistics were calculated for the scores of the two test sessions. Independent samples *t*-tests were used to analyse differences in age and exercise capacity between patients and the healthy subjects group I and group II. Test–retest reliability was analyzed by means of a paired *t*-test and intra class correlation coefficient (ICC, one way random model). For the reference range of the differences between the two measurements of the LBM-based Åstrand bicycle test the limits of agreement (LOA) were calculated

as $\pm 1.96 * SD_{difference}$. The repeatability coefficient (RC) was calculated, from the results of one-way analyse of variance (ANOVA), with the subject as factor [19–21]. The within subjects standard deviation (S_w) was estimated as the square root of the residual mean square. The RC was calculated as $\pm 1.96 * \sqrt{2}S_w$ [20]. The RC represents the limits within which it is expected the differences between two measurements by the same method will lie [20]. In case of significant differences between the first and second test in the paired *t*-test, the ICC, the LOA, and RC were not calculated because test results differ systematically [20, 21]. An ICC $\geq$ 0.75 was considered an acceptable level of reliability [21, 22]. No criteria for interpretation of the LOA and RC are available. However, a smaller LOA indicate more stability over time because it indicates that the natural variation is small [14, 23]. Plots were made of the individual difference between sessions against the individual mean of the two sessions, to analyze whether the magnitude of the difference was related to the mean performance [21]. A funnel shape indicates that the magnitude of the difference is related to the mean performance [21]. LOA was expressed as a percentage of the mean of two measurements [14]. The validity of the LBM-based Åstrand bicycle test was evaluated by means of a paired *t*-test and an ICC between the LBM-based VO_2 max and the VO_2 max based on the maximal bicycle test. Agreement in between the two methods was calculated as the LOA. LOA were also expressed as a percentage of the mean of two measurements [20]. Data analyses were performed using the Statistical Package for the Social Sciences (SPSS 14.0).

Results

The outcomes of the two LBM-based Åstrand bicycle tests did not differ significantly in patients or in healthy subjects (Table 2A). The Bland and Altman plots did not show funnel shapes (Fig. 1), indicating that the difference between the two sessions is not related to the average performances on the two sessions. In figure one an outlier was identified. This outlier influenced the outcomes considerable, therefore outcomes are presented in two ways, one including the outlier (Table 2A) and one excluding this outlier (Table 2B) [20]. The ICC's between the first and the second LBM-based bicycle test results ranged in patients from 0.91 to 0.94 including the outlier and in healthy subjects from 0.97 to 0.98 (Table 2A) and excluding the outlier ranged in patients from 0.96 to 0.98 (Table 2B). The LOA and RC in patients including the outlier are considerably wider than in healthy controls. When the outlier was exclude LOA and RC of patients and healthy subjects were similar (Figs. 2 and 3). The LOA percentages of the mean of two measurements ranged in patients include the outlier from 32.0 to 32.8% (Table 2A) and excluding the outlier from the analysis from 13.8 to 16.9% (Table 2B). In healthy subjects the LOA percentages of the mean of two

Table 2 (A) Results of the test–retest reliability for patients and healthy subjects and (B) Results of the test–retest reliability for patients, without the outlier ($n = 18$) in the first and second measurement sessions of the LBM-based Åstrand bicycle test

	T1(SD)	T2(SD)	Mean of two measurements(SD)	ΔT(SD)	ICC(95% CI)	RC	LOA	LOA% of the mean of two measurements
(A)								
Patients (n = 19)								
ml/kg LBM * min^{-1}	51.1(16.5)	51.1(11.4)	51.1(13.5)	0.02(8.4)	0.91(0.76–0.97)	±16.1	±16.6	32.2
ml/kg BM * min^{-1}	38.6(15.1)	37.9(11.6)	38.3(13.0)	0.63(6.4)	0.94(0.85–0.98)	±12.3	±12.6	32.8
l/min	2.84(1.1)	2.78(0.8)	2.81(0.9)	0.05(0.4)	0.91(0.76–0.97)	±0.9	±0.9	32.0
Healthy subjects (n = 20)								
ml/kg LBM * min^{-1}	58.7(12.7)	58.1(13.1)	58.3(12.8)	0.63(4.1)	0.97(0.94–0.99)	±8.0	±8.0	13.7
ml/kg BM * min^{-1}	48.8(11.2)	48.4(12.2)	48.6(11.6)	0.50(3.6)	0.98(0.94–0.99)	±7.1	±7.1	14.6
l/min	3.55(0.9)	3.51(1.1)	3.53(1.0)	0.04(0.2)	0.97(0.94–0.99)	±0.5	±0.5	14.1
(B)								
Patients (n = 18)								
ml/kg LBM * min^{-1}	48.7(13.9)	50.6(11.5)	49.7(12.6)	−1.71(4.3)	0.96(0.91–0.99)	±8.4	±8.4	16.9
ml/kg BM * min^{-1}	36.4(12.5)	37.2(11.5)	36.9(11.9)	−0.75(2.6)	0.98(0.97–1.00)	±5.1	±5.1	13.8
l/min	2.69(0.9)	2.73(0.8)	2.71(0.9)	−0.05(0.2)	0.98(0.95–0.99)	±0.4	±0.4	14.7

There were no significant differences between the first and second test in the paired *t*-test of the LBM-based Åstrand bicycle test

T1: mean value of the first assessment, T2: mean value of the second assessment (SD): standard deviation, ΔT: difference between T1 and T2

ICC: Intra Class Correlation, CI: Confidence of Interval, RC: repeatability coefficient, LOA: Limits of agreement, LBM: Lean Body Mass, BM: Body Mass

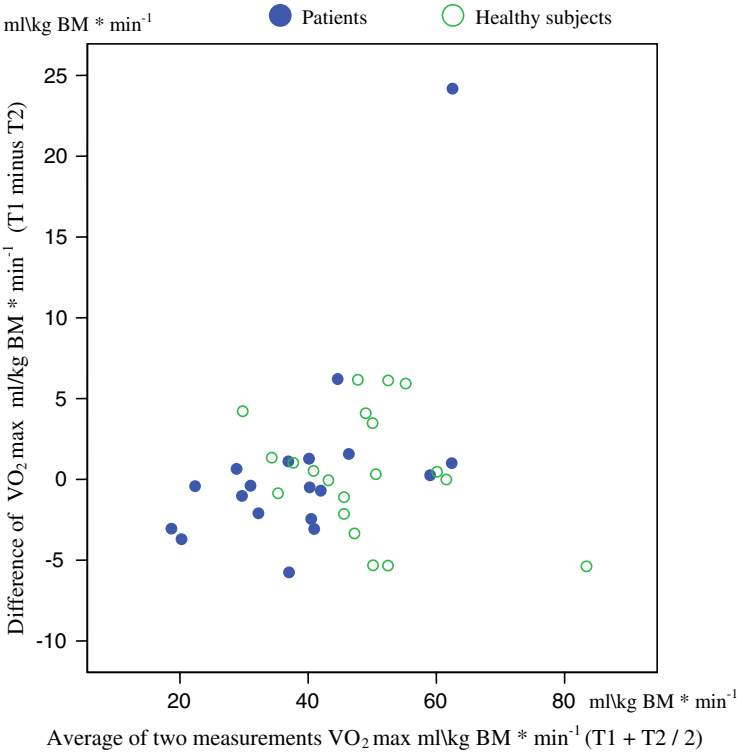

Fig. 1 LBM-based Åstrand bicycle test in patients and healthy subjects. Difference between T1 and T2 scatter plotted against T1 + T2 divided by 2

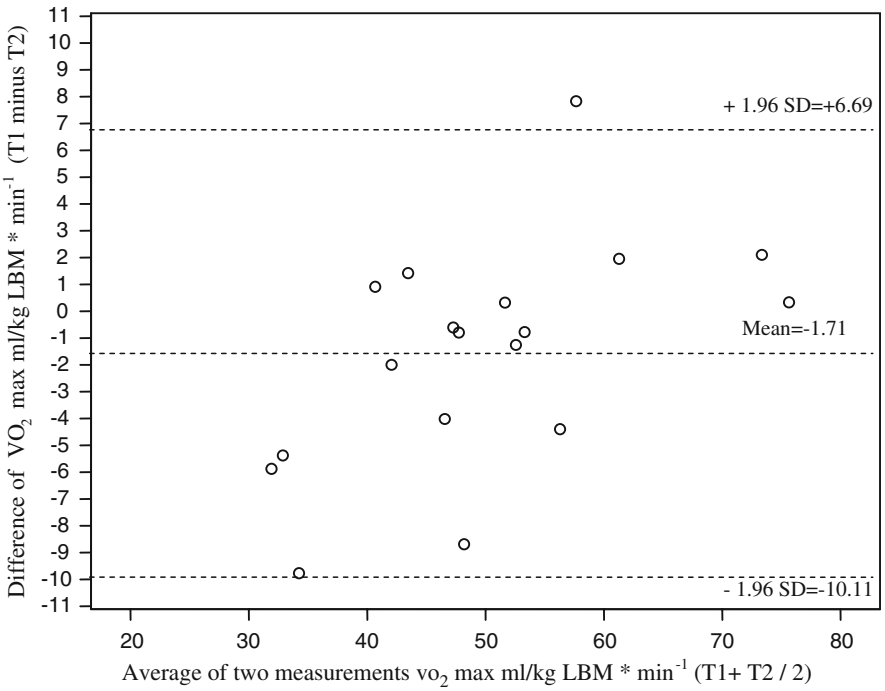

Fig. 2 Bland and Altman scatter plot LBM-based Åstrand bicycle test in patients. VO_2 max ml/kg LBM * min^{-1} of the two test for the test–retest reliability with exclusion of the outlier

measurements ranged 13.7–14.6% (Table 2A). Only one patient (5%) of our study population stopped the test prematurely due to fatigue and pain.

The patients are significantly older than the healthy subjects of group I ($P < 0.001$). The aerobic capacity in the healthy subject group I are significantly higher in ml/kg LBM * min^{-1} ($P < 0.045$) ml/kg BM * min^{-1} ($P < 0.005$) and L/min ($P < 0.014$) than in patients.

The outcomes of the LBM-based Åstrand bicycle test and the maximal bicycle test did not differ significantly

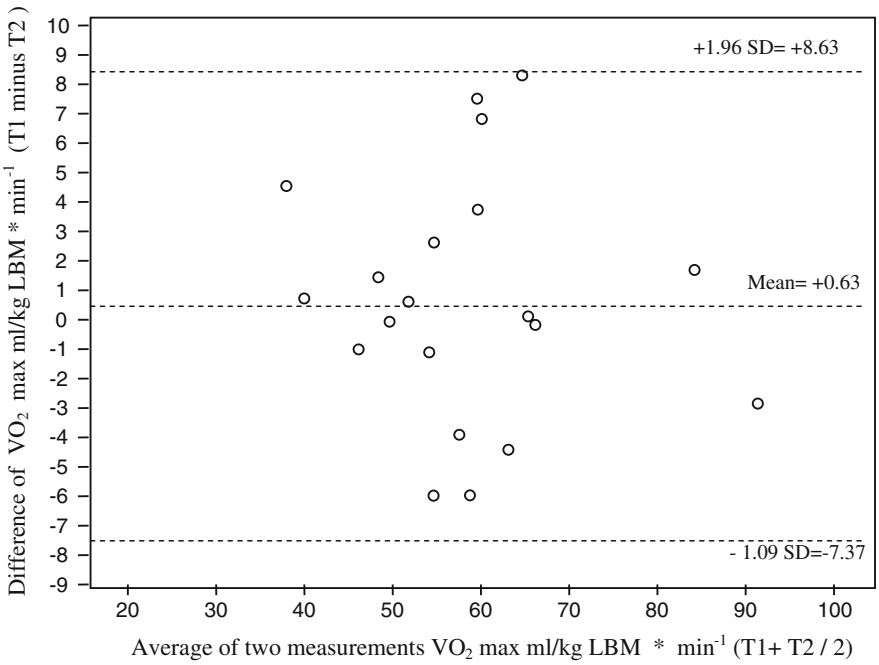

Fig. 3 Bland and Altman scatter plot LBM-based Åstrand bicycle test in healthy subjects. VO_2 max ml/kg LBM * min^{-1} of the two test for the test–retest reliability

Table 3 Results for healthy subjects in the first measurement session of the LBM-based Åstrand bicycle test and the second measurement session of the maximal bicycle test

Healthy subjects ($n = 19$)	LBM-based Åstrand test T1 (SD)	Maximal bicycle test T2 (SD)	Mean of two measurements (SD)	ΔT (SD)	ICC (95% CI)	95% LOA	LOA% of the mean of two measurements
ml/kg LBM * min^{-1}	66.2(16.7)	63.9(12.7)	65.3(14.1)	2.72(9.6)	0.88(0.70–0.95)	±18.8	28.7
ml/kg BM * min^{-1}	53.5(13.8)	51.8(13.4)	52.7(13.2)	1.63(6.9)	0.93(0.82–0.97)	±13.5	25.6
l/min	3.80(1.1)	3.60(1.1)	3.70(1.1)	0.20(0.5)	0.95(0.88–0.98)	±0.9	24.3

There were no significant differences between the LBM-based Åstrand bicycle test and the maximal bicycle test

T1: mean value of the first assessment, T2: mean value of the second assessment (SD): standard deviation, ΔT: difference between T1 and T2

ICC: Intra Class Correlation, CI: Confidence of Interval, LOA: Limits of Agreement, LBM: Lean Body Mass, BM: Body Mass

(Table 3). The ICC's between the LBM-based Åstrand bicycle test and the maximal bicycle test ranged from 0.88 to 0.95 (Table 3). The healthy subjects of group II are significantly younger ($P < 0.000$) then the patients. The aerobic capacity in the healthy subjects of group II are significantly higher in ml/kg.LBM * min^{-1} ($P < 0.008$) ml/kg.BM * min^{-1} ($P < 0.003$) and L/min ($P < 0.012$) than in patients.

Discussion

The results of our study show that the LBM-based Åstrand bicycle test is reliable, valid, and feasible, although there is considerable test–retest variation. The ICC is ≥0.91 for all comparisons of the LBM-based Åstrand bicycle tests in patients and healthy subjects. This shows that the LBM-based Åstrand bicycle test is reliable in patients as well as in healthy subjects. However, the ICC expresses how well observations are likely to classify a patient consistently relative to other patients [14, 23]. The ICC value provides no indication of the magnitude of the variation between two observations (within-patient variance or 'noise') [24]. To quantify the magnitude of such variation the LOA and RC were calculated [20, 21]. The LOA reflects the average differences between two measurements and the RC reflects the differences with the estimated true value of an individual. The RC is a very similar analysis to the limits of agreement approach and can also be applied to quantify the repeatability of a method from replicated measurements obtained by the same method [20]. A smaller RC indicates less intra individual variation. Clinically the LOA indicates that, in an individual patient with non-specific CLBP, the changes due to treatment should exceed the 95% LOA

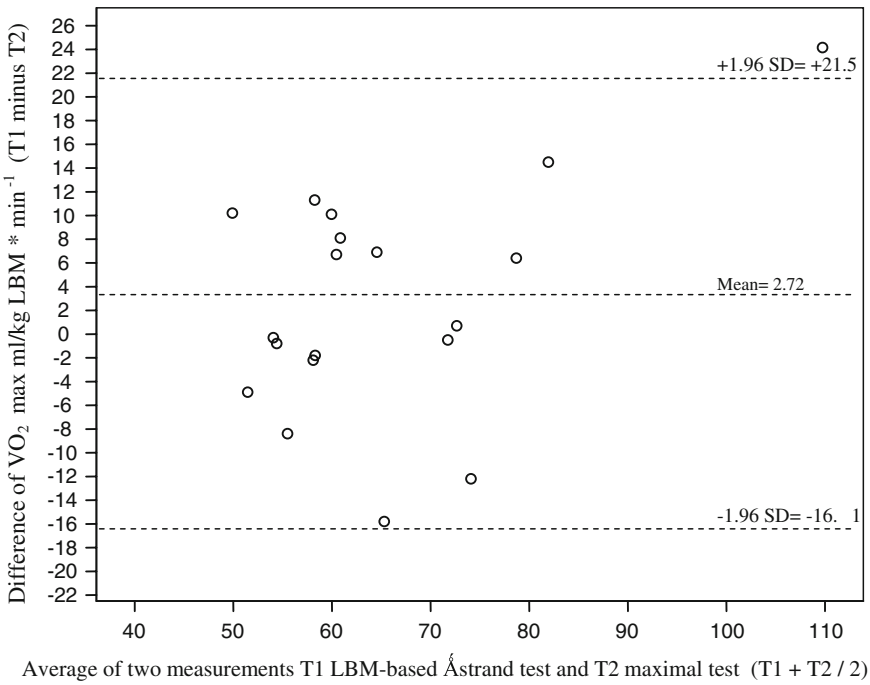

Fig. 4 Bland and Altman scatter plot LBM-based Åstrand bicycle test and maximal test in healthy subjects. VO_2 max ml/kg LBM * min^{-1} of the two test for the validity

before permitting the conclusion that a true change, increase or decrease, in exercise capacity has occurred [14, 23]. Despite the very good reliability, the LOA and RCs are considerably wider than in healthy subjects. These findings indicate that true change in an individual patient can only be detected when at least 33% of changes include the outlier (Table 2A) and 17% of changes exclude the outlier (Table 2B) of the mean of two measurements. However, excluding the outlier the result of LOA and RC were similar for patients and healthy subjects. The mean outcome of the two LBM-based Åstrand bicycle test was significantly lower for patients compared to the control group(s). Similar results were found previously in which patients with CLBP had a significant lower VO_2 max than the matched healthy referents [7]. However, it is conceivable that in our study the significant difference in age between patients and healthy subjects contributed to the differences in aerobic capacity. The severity and duration of back complaints of our patients is similar to that of patients in other studies. The median duration of complaints of our study population of 68 months (range 8–180 months) is similar with those of other studies of 48 months (range 16–120 months) [10] and 62 months (range 3–396 months) [7]. The mean RMDQ score of our patients of 10.2 (5.3) was also similar to those of other studies, 10.17 (6.22) [25] and 14.2 (3.9) [7]. Duration of complaints of 68 months can be considered as a chronic condition. A RMDQ score of 10 is considered as a mean disability score in patients with CLBP [26].

The validity of the LBM-based Åstrand bicycle test was determined using the ICC coefficient with the maximal bicycle test. The ICC's between the LBM-based Åstrand bicycle test and the maximal bicycle test were ≥0.88 (Table 3). Based on these results, we can conclude that the LBM-based Åstrand bicycle test is valid in healthy subjects. However, it should be noted that the LOA percentage of the mean of two measurements were substantial, ranging from 24 to 29%. The Bland and Altman scatter plot (Fig. 4) shows that the range is +21.5 to −16.1 VO_2 max ml/kg LBM * min^{-1}. Further research is necessary to identify the sources of this variation. Although there is a significant difference in age between patients and healthy subjects in our study, it is conceivable that the LBM-based Åstrand bicycle test is also valid in non-specific CLBP patients. This is supported by Macsween [27] and Cink [28] that extrapolation of submaximal data using the Åstrand age correction factors is valid.

The feasibility of the LBM-based Åstrand bicycle test in our patients is good, with only one patient (5%) stopping due to fatigue or pain. In a previous study, it was reported that 12% of 84 patients with non-specific CLBP stopped the LBM-based Åstrand bicycle test prematurely due to pain or fatigue [7]. Feasibility of the LBM-based Åstrand bicycle test is higher than the conventional Åstrand bicycle test in patients with non-specific CLBP [5, 7]. Compared to a submaximal bicycle ergometer test, 33% of 504 patients with chronic lumbar spinal disorder were unable to finish the test [11]. In a Posthoc analysis feasibility of the LBM-based Åstrand bicycle test in our study and that from Smeets [7] appeared significantly better compared to the bicycle test from Protas [11] (chi-square test respectively $P < 0.004$ and $P < 0.001$). Moreover, previous studies

showed that of patients who underwent a symptom-limited modified treadmill test, 25 (50%) of 50 CLBP patients stopped because of pain and 21 (42%) stopped because of fatigue. The remaining 8% stopped because of test termination [1]. The outcomes of our study support the findings that the LBM-based Åstrand bicycle test increases significantly the number of patients that can finish the test. These findings indicate that the test has considerable clinical advantages above the normal Åstrand test. There are several limitations in our study. A limitation of our study was that our patients may have been a selection of non-specific CLBP patients because they were motivated to do the exercise test twice therefore the feasibility in our study is likely to be higher than in general non-specific CLBP population. Additionally our results are based on the small group of non-specific CLBP patients and sample variation may have influenced our results.

In conclusion, this study shows that the LBM-based Åstrand bicycle test is reliable, valid, and feasible. However, a substantial amount of variation should be taken into account in patients when interpreting the test results clinically.

Acknowledgements We thank all participating patients, healthy subjects, rehabilitation physicians and physical therapists for their time and effort. We thank Roy E. Stewart for helping with data analyzing.

References

1. Wittink H, Hoskins Michel T, Wagner A, Sukiennik A, et al. Deconditioning in patients with chronic low back pain. Fact or fiction? Spine. 2000;25:2221–8. doi:10.1097/00007632-200009010-00013.
2. Brox JI, Storheim K, Holm I, Friis A, Reikeras O. Disability, pain, psychological factors, physical performance in healthy controls, patients with sub-acute and chronic low back pain: a case control study. J Rehabil Med. 2005;37:95–9. doi:10.1080/16501970410017738.
3. Hurri H, Mellin G, Korhonen O, Harjula R, et al. Aerobic capacity among chronic low back pain patients. J Spinal Disord. 1991;4:34–8.
4. Nielens H, Plaghki L. Perception of pain exertion during exercise on a cycle ergometer on chronic pain patients. Clin J Pain. 1994;10:204–9. doi:10.1097/00002508-199409000-00006.
5. Verbunt JA, Seelen HA, Vlaeyen JWS, van der Heijden GJ, et al. Fear of injury and physical deconditioning in patients with chronic low back pain. Arch Phys Med Rehabil. 2003;84:1227–32. doi:10.1016/S0003-9993(03)00132-1.
6. Verbunt JA, Seelen HA, Vlaeyen JWS, van der Heijden GJ, et al. Disuse and deconditioning in chronic low back: concepts and hypothesis on contributing mechanisms. Eur J Pain (Lond, Engl). 2003;7:9–21. doi:10.1016/S1090-3801(02)00071-X.
7. Smeets RJEM, Wittink H, Hidding A, Knottnerus A. Do patients with chronic low back pain have a lower level of aerobic fitness than healthy controls? Spine. 2006;31:90–7. doi:10.1097/01.brs.0000192641.22003.83.
8. Koho P, Sarita A, Watson P, Hurri H. Assessment of chronic pain behaviour: Reliability of the method and its relationship with perceived disability, physical impairment and function. J Rehabil Med. 2001;33:128–32. doi:10.1080/165019701750165970.
9. Hodselmans AP, Jaegers SM, Göeken LNH. Short-term outcomes of a back school programme for chronic low back pain. Arch Phys Med Rehabil. 2001;82:1099–105. doi:10.1053/apmr.2001.23899.
10. Keller A, Hellesnes J, Brox JI. Reliability of the isokinetic trunk extensor test, Biering-Sorensen test and Åstrand bicycle test: assessment of intra class correlation coefficient and critical difference in patients with chronic low back pain and healthy individuals. Spine. 2001;26:771–7. doi:10.1097/00007632-200104010-00017.
11. Protas E, Mayer T, Dersh J, Keely J, Gatchel R, McGeary D. Relevance of aerobic capacity measurements in the treatment of chronic work-related spinal disorders. Spine. 2004;29:2158–66. doi:10.1097/01.brs.0000141154.38841.05.
12. Wittink H, Rogers W, Gascon C, Sukiennik A, et al. Relative contribution of mental health and exercise related pain increment to treadmill test intolerance in patients with chronic low back pain. Spine. 2001;26:2368–74. doi:10.1097/00007632-200111010-00014.
13. Greenleaf JE. Intensive exercise training during bed rest attenuates deconditioning. Med Sci Sports Exerc. 1997;29:207–15. doi:10.1097/00005768-199702000-00007.
14. Hodselmans AP, Dijkstra PU, van der Schans CP, Geertzen JHB. Test–retest reliability of psychophysical lift capacity in patients with non-specific chronic low back pain and healthy subjects. J Rehabil Med. 2007;39:133–7. doi:10.2340/16501977-0014.
15. Durnin JVGA, Womersly J. Body fat assessed from total body density and its estimation from skinfold thickness: measurements on 481 men and women aged 16–72 years. Br J Nutr. 1974;32:77–97. doi:10.1079/BJN19740060.
16. Binkhorst RA. Guidelines for the ergometry in sport medical advise center NISGZ. The Netherlands: Oosterbeek; 1986.
17. Åstrand PO, Rodahl K. Textbook of work physiology, physiological bases of exercise. New York: McGraw-Hill; 1986.
18. Fox S. Physical activity and the prevention of coronary heart disease. Ann Clin Res. 1971;3:404–32.
19. Bland JM, Altman DG. Measurement error. BMJ (Clinical Research Ed). 1996;313:744.
20. Bland JM, Altman DG. Measuring agreement in method comparison studies. Stat Methods Med Res. 1999;8:135–60. doi:10.1191/096228099673819272.
21. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet. 1986;1:307–10.
22. Lee J, Koh D, Ong CN. Statistical evaluation of agreement between two methods for measuring quantitative variables. Comput Biol Med. 1989;19:61–70. doi:10.1016/0010-4825(89)90036-X.
23. Brouwer S. Disability in chronic low back pain. Psychometric properties of ADL- and work-related instruments [dissertation]. Groningen: University Groningen; 2004.
24. Evans WJ, Cayten CG, Green PA. Determining the generalizability of rating scales in clinical settings. Med Care. 1981;19:1211–20. doi:10.1097/00005650-198112000-00004.
25. Cunha IT, Simmonds MJ, Protas EJ, Jones S. Back pain, physical function and estimates of aerobic capacity. What are the relations among methods and measures? Am J Phys Med Rehabil. 2002;81:913–20. doi:10.1097/00002060-200211000-00008.
26. Roland M, Morris R. A study of the natural history of low-back pain. Part II. development of guidelines for trials of treatment in primary care. Spine. 1983;8:145–50. doi:10.1097/00007632-198303000-00005.
27. Macsween A. The reliability and validity of the Astrand nomogram and linear extrapolation for derving VO_2 max from submaximal exercise data. J Sports Med Phys Fit. 2001;41:312–7.
28. Cink RE, Thomas TR. Validity-rhyming nomogram for predicting maximal oxygen intake. Br J Sports Med. 1981;15:182–5.